Appendix

We let Y_a and M_a denote respectively the values of the outcome and mediator that would have been observed had the exposure A been set to level a. We let Y_{am} denote the value of the outcome that would have been observed had the exposure, A, and mediator, M, been set to levels a and m, respectively.

The average controlled direct effect comparing exposure level a to a^* and fixing the mediator to level m is defined by $CDE_{a,a^*}(m) = E[Y_{am} - Y_{a^*m}]$. The average natural direct effect is then defined by $NDE_{a,a^*}(a^*) = E[Y_{aM_{a^*}} - Y_{a^*M_{a^*}}]$. The average natural indirect effect can be defined as $NIE_{a,a^*}(a) = E[Y_{aM_a} - Y_{aM_{a^*}}]$, which compares the effect of the mediator at levels M_a and M_{a^*} on the outcome when exposure A is set to a. Controlled direct effects and natural direct and indirect effects within strata of $C = c$ are then defined by: $CDE_{a,a^*|c}(m) = E[Y_{am} - Y_{a^*m}|c]$, $NDE_{a,a^*|c}(a^*) = E[Y_{aM_{a^*}} - Y_{a^*M_{a^*}}|c]$ and $NIE_{a,a^*|c}(a) = E[Y_{aM_a} - Y_{aM_{a^*}}|c]$ respectively.

For a dichotomous outcome the total effect on the odds ratio scale conditional on $C = c$ is given by $OR^{TE}_{a,a^*|c} = \frac{P(Y_a=1|c)/\{1-P(Y_a=1|c)\}}{P(Y_{a^*}=1|c)/\{1-P(Y_{a^*}=1|c)\}}$. The controlled direct effect on the odds ratio scale is given by $OR^{CDE}_{a,a^*|c}(m) = \frac{P(Y_{am}=1|c)/\{1-P(Y_{am}=1|c)\}}{P(Y_{a^*m}=1|c)/\{1-P(Y_{a^*m}=1|c)\}}$. The natural direct effect on the odds ratio scale conditional on $C = c$ is given by $OR^{NDE}_{a,a^*|c}(a^*) = \frac{P(Y_{aM_{a^*}}=1|c)/\{1-P(Y_{aM_{a^*}}=1|c)\}}{P(Y_{a^*M_{a^*}}=1|c)/\{1-P(Y_{a^*M_{a^*}}=1|c)\}}$. The natural indirect effect on the odds ratio scale conditional on $C = c$ is given by $OR^{NIE}_{a,a^*|c}(a) = \frac{P(Y_{aM_a}=1|c)/\{1-P(Y_{aM_a}=1|c)\}}{P(Y_{aM_{a^*}}=1|c)/\{1-P(Y_{aM_{a^*}}=1|c)\}}$.

As discussed in the text, identification assumptions (i)-(iv) will suffice to identify these direct and indirect effects. If we let $X \perp Y|Z$ denote that X is independent of Y conditional on Z then these four identification assumptions can be expressed formally in terms of counterfactual independence as (i) $Y_{am} \perp A|C$, (ii) $Y_{am} \perp M|\{A, C\}$, (iii) $M_a \perp A|C$, and (iv) $Y_{am} \perp M_{a^*}|C$. Assumptions (i) and (ii) suffice to identify controlled direct effects; assumptions (i)-(iv) suffice to identify natural direct and indirect effects (Pearl, 2001; VanderWeele and Vansteelandt, 2009). The intuitive interpretation of these assuptions as described in the text follows from the theory of causal diagrams (Pearl, 2001). Alternative identification assumptions have also been proposed (Imai 2010a; Hafeman and VanderWeele, 2011). However, it has been shown that the intuitive graphical interpretation of these alternative assumptions are in fact equivalent (Shpitser and VanderWeele, 2011). Technical examples can be constructed where one set of identifiation assumptions holds and another does not, but on a causal diagram corresponding to a set of non-parametric structural equations, whenever one set of the assumptions among those in VanderWeele and Vansteelandt (2009), Imai (2010a), and Hafeman and VanderWeele (2011) holds, the others will also.

1 Continuous Mediator and Outcome

Effects using regression

Suppose that both the mediator and the outcome are continuous and that the following models fit the observed data:

$$E(M|A=a, C=c) = \beta_0 + \beta_1 a + \beta_2' c \tag{1}$$

$$E(Y|A=a, M=m, C=c) = \theta_0 + \theta_1 a + \theta_2 m + \theta_3 a * m + \theta_4' c \tag{2}$$

If the covariates C satisfied the no-unmeasured confounding assumptions (i)-(iv) above, then the average controlled direct effect and the average natural direct and indirect effects were derived by VanderWeele and Vansteelandt, 2009.

In particular, if the regression models (1) and (2) are correctly specified and assumptions of no unmeasured confounding of exposure-outcome relationship (i) and no unmeasured confounding of the mediator-outcome relationship (ii) hold, then we could compute the controlled direct effect as follows:

$CDE = E[Y_{am} - Y_{a^*m}|C=c]$

$= E[Y|C = c, A = a, M = m] - E[Y|C = c, A = a^*, M = m]$

$= (\theta_0 + \theta_1 a + \theta_2 m + \theta_3 am + \theta_4' c) - (\theta_0 + \theta_1 a^* + \theta_2 m + \theta_3 a^* m + \theta_4' c)$

$= (\theta_1 a + \theta_3 am - \theta_1 a^* - \theta_3 a^* m)$

$= \theta_1(a - a^*) + \theta_3 m(a - a^*).$

If the regression models (1) and (2) are correctly specified and assumptions (*i*) and (*ii*) together with two additional assumptions of (*iii*) no unmeasured confounding of the exposure-mediator relationship and (*iv*) that there is no mediator-outcome confounder that is affected by the exposure hold, then we could compute the natural direct effects by:

$NDE = E[Y_{aM_{a^*}} - Y_{a^*M_{a^*}}|C = c]$

$= \sum_m \{E[Y|C = c, A = a, M = m] - E[Y|C = c, A = a^*, M = m]\} \times P(M = m|C = c, A = a^*)$

$= \sum_m \{(\theta_0 + \theta_1 a + \theta_2 m + \theta_3 am + \theta_4' c) - (\theta_0 + \theta_1 a^* + \theta_2 m + \theta_3 a^* m + \theta_4' c)\} \times P(M = m|C = c, A = a^*)$

$= \sum_m \{(\theta_1 a + \theta_2 m + \theta_3 am) - (\theta_1 a^* + \theta_2 m + \theta_3 a^* m)\} \times P(M = m|C = c, A = a^*)$

$= \{(\theta_1 a + \theta_2 E[M|A = a^*, C = c] + \theta_3 a E[M|A = a^*, C = c]) - (\theta_1 a^* + \theta_2 E[M|A = a^*, C = c] + \theta_3 a^* E[M|A = a^*, C = c])\}$

$= \{(\theta_1 a + \theta_2(\beta_0 + \beta_1 a^* + \beta_2' c) + \theta_3 a(\beta_0 + \beta_1 a^* + \beta_2' c) - (\theta_1 a^* + \theta_2(\beta_0 + \beta_1 a^* + \beta_2' c) + \theta_3 a^*(\beta_0 + \beta_1 a^* + \beta_2' c))\}$

$= \{\theta_1 a + \theta_3 a(\beta_0 + \beta_1 a^* + \beta_2' c) - (\theta_1 a^* + \theta_3 a^*(\beta_0 + \beta_1 a^* + \beta_2' c))\}$

$= (\theta_1 + \theta_3\beta_0 + \theta_3\beta_1 a^* + \theta_3\beta_2' c)(a - a^*).$

Moreover under the same assumptions we can compute the natural indirect effects by:

$NIE = E[Y_{aM_a} - Y_{aM_{a^*}}|C = c]$

$= \sum_m E[Y|C = c, A = a, M = m] \times P(M = m|C = c, A = a) - \sum_m E[Y|C = c, A = a, M =$

$m] \times P(M=m|C=c, A=a^*)$

$= \sum_m(\theta_0 + \theta_1 a + \theta_2 m + \theta_3 am + \theta_4' c) \times P(M=m|C=c, A=a) - \sum_m(\theta_0 + \theta_1 a + \theta_2 m + \theta_3 am + \theta_4' c) \times P(M=m|C=c, A=a^*)$

$= (\theta_0 + \theta_1 a + \theta_2 E[M|A=a, C=c] + \theta_3 a E[M|A=a, C=c] + \theta_4' c) - (\theta_0 + \theta_1 a + \theta_2 E[M|A=a^*, C=c] + \theta_3 a^* E[M|A=a^*, C=c] + \theta_4' c)$

$= (\theta_1 a + \theta_2(\beta_0 + \beta_1 a + \beta_2' c) + \theta_3 a(\beta_0 + \beta_1 a + \beta_2' c)) - (\theta_1 a^* + \theta_2(\beta_0 + \beta_1 a^* + \beta_2' c) + \theta_3 a^*(\beta_0 + \beta_1 a^* + \beta_2' c))$

$= (\theta_2\beta_1 + \theta_3\beta_1 a)(a - a^*).$

If the regression models (1) and (2) are correctly specified and assumptions (i) and (ii) hold, then we could compute the total effect by:

$TE = E[Y_a - Y_{a^*}|C=c]$

$= E[Y_{a,M(a^*)} - Y_{a^*,M(a^*)}|C=c] + E[Y_{a,M(a)} - Y_{a^*,M(a^*)}|C=c]$

$= (\theta_1 + \theta_3\beta_0 + \theta_3\beta_1 a^* + \theta_3\beta_2' c + \theta_2\beta_1 + \theta_3\beta_1 a)(a - a^*).$

Finally if the regression models (1) and (2) are correctly specified and assumptions (i)-(iv) hold then we could compute the proportion mediated by:

$PM = \frac{E[Y_{aM_a} - Y_{a^*M_{a^*}}|C=c]}{E[Y_a - Y_{a^*}|C=c]}$

$= \frac{\theta_2\beta_1 + \theta_3\beta_1 a}{\theta_1 + \theta_3\beta_0 + \theta_3\beta_1 a^* + \theta_3\beta_2' c + \theta_2\beta_1 + \theta_3\beta_1 a}.$

Standard errors

Suppose that model (1) and (2) have been fit using standard linear regression software and that the resulting estimates $\hat{\beta}$ of $\beta = (\beta_0, \beta_1, \beta_2')'$ and $\hat{\theta}$ of $\theta = (\theta_0, \theta_1, \theta_2, \theta_3, \theta_4')'$ have covariance matrices Σ_β and Σ_θ. Then the covariance matrix of $(\hat{\beta}', \hat{\theta}')$ is

$$\Sigma = \begin{bmatrix} \Sigma_\beta & 0 \\ 0 & \Sigma_\theta \end{bmatrix}$$

Standard errors of the controlled and natural direct and indirect effects can be obtained (using the delta method) as

$$\sqrt{\Gamma\Sigma\Gamma'}|a - a^*|$$

with $\Gamma = (0, 0, 0', 0, 1, 0, m, 0')$ for the controlled direct effect, $\Gamma = (\theta_3, \theta_3 a^*, \theta_3 c', 0, 1, \beta_0 + \beta_1 a^* + \beta_2' c, 0')$ for the pure natural direct effect (same expression holds for the total natural direct effect upon substituting a and a^*), $\Gamma = (0, \theta_2 + \theta_3 a, 0', 0, 0, \beta_1, \beta_1 a, 0')$ for the total natural indirect effect (the same expression holds for the pure natural indirect effect upon substituting a and a^*), $\Gamma = (\theta_3, \theta_3(a + a^*) + \theta_2, \theta_3 c', 0, 1, \beta_1, \beta_0 + \beta_1(a + a^*) + \beta_2' c, 0')$ for the total effect and for the proportion mediated $\Gamma = (d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8)$ where

$$d_1 = -\theta_3 \frac{\theta_2\beta_1 + \theta_3\beta_1 a}{(\theta_1 + \theta_3\beta_0 + \theta_3\beta_1 a^* + \theta_3\beta_2' c + \theta_2\beta_1 + \theta_3\beta_1 a)^2}$$

$$d_2 = \frac{(\theta_2 + \theta_3 a)(-(\theta_2\beta_1 + \theta_3\beta_1 a) + (\theta_1 + \theta_3\beta_0 + \theta_3\beta_1 a^* + \theta_3\beta_2' c + \theta_2\beta_1 + \theta_3\beta_1 a)) - \theta_3 a^*}{(\theta_1 + \theta_3\beta_0 + \theta_3\beta_1 a^* + \theta_3\beta_2' c + \theta_2\beta_1 + \theta_3\beta_1 a)^2}$$

$$d_3 = -\frac{\theta_3 c'(\theta_2\beta_1 + \theta_3\beta_1 a)}{(\theta_1 + \theta_3\beta_0 + \theta_3\beta_1 a^* + \theta_3\beta_2' c + \theta_2\beta_1 + \theta_3\beta_1 a)^2}$$

$$d_4 = 0$$

$$d_5 = -\frac{\theta_2\beta_1 + \theta_3\beta_1 a}{(\theta_1 + \theta_3\beta_0 + \theta_3\beta_1 a^* + \theta_3\beta_2' c + \theta_2\beta_1 + \theta_3\beta_1 a)^2}$$

$$d_6 = \frac{\beta_1(-(\theta_2\beta_1 + \theta_3\beta_1 a) + (\theta_1 + \theta_3\beta_0 + \theta_3\beta_1 a^* + \theta_3\beta_2^{'}c + \theta_2\beta_1 + \theta_3\beta_1 a))}{(\theta_1 + \theta_3\beta_0 + \theta_3\beta_1 a^* + \theta_3\beta_2^{'}c + \theta_2\beta_1 + \theta_3\beta_1 a)^2}$$

$$d_7 = \frac{\beta_1 a(\theta_1 + \theta_3\beta_0 + \theta_3\beta_1 a^* + \theta_3\beta_2^{'}c + \theta_2\beta_1 + \theta_3\beta_1 a) - (\beta_0 + \beta_1(a + a^*) + \beta_2^{'}c)(\theta_2\beta_1 + \theta_3\beta_1 a)}{(\theta_1 + \theta_3\beta_0 + \theta_3\beta_1 a^* + \theta_3\beta_2^{'}c + \theta_2\beta_1 + \theta_3\beta_1 a)^2}$$

$$d_8 = 0^{'}.$$

2 Continuous Mediator and Binary Outcome

Effects using regression

Suppose that the mediator is continuous and the outcome is binary and is rare. Suppose that the following models fit the observed data:

$$E(M|A = a, C = c) = \beta_0 + \beta_1 a + \beta_2^{'}c \quad (3)$$

$$logit\{P(Y = 1|A = a, M = m, C = c)\} = \theta_0 + \theta_1 a + \theta_2 m + \theta_3 a * m + \theta_4^{'}c \quad (4)$$

and that the error term in the regression model for M is normally distributed with mean 0 and variance σ^2. If the regression models (3) and (4) are correctly specified and assumptions (i) and (ii) hold then the conditional controlled direct effect on the odds ratio scale would be given by (VanderWeele and Vansteelandt, 2010):

$OR^{CDE} = \frac{P(Y_{am}=1|c)/(1-P(Y_{am}=1|c))}{P(Y_{a^*m}=1|c)/(1-P(Y_{a^*m}=1|c))}$
$= \frac{P(Y=1|a,m,c)/(1-P(Y=1|a,m,c))}{P(Y=1|a^*,m,c)/(1-P(Y=1|a^*m,c))}$

$= \frac{exp[\theta_0+\theta_1 a+\theta_2 m+\theta_3 am+\theta_4^{'} c]}{exp[\theta_0+\theta_1 a^*+\theta_2 m+\theta_3 a^* m+\theta_4^{'} c]}$
$= exp[(\theta_1 + \theta_3 m)(a - a^*)].$

If the regression models (3) and (4) are correctly specified and assumptions (i)-(iv) hold, the outcome Y is rare, and the error term for linear regression model (1) is normally distributed and has constant variance σ^2, then we could compute the natural direct effects by:

$OR^{NDE} = exp[log\{\frac{P(Y_{aM_{a^*}}=1|c)/(1-P(Y_{aM_{a^*}}=1|c))}{P(Y_{a^*M_{a^*}}=1|c)/(1-P(Y_{a^*M_{a^*}}=1|c))}\}]$
$= exp[logit\{P(Y_{aM_{a^*}} = 1|c)\} - logit\{P(Y_{a^*M_{a^*}} = 1|c)\}]$
$\sim exp[\theta_0 + \theta_1 a + \theta_4^{'} c + (\theta_2 + \theta_3 a)(\beta_0 + \beta_1 a^* + \beta_2^{'} c) + \frac{1}{2}(\theta_2 + \theta_3 a)^2\sigma^2 - \{\theta_0 + \theta_1 a^* + \theta_4^{'} c + (\theta_2 + \theta_3 a)(\beta_0 + \beta_1 a^* + \beta_2^{'} c) + \frac{1}{2}(\theta_2 + \theta_3 a^*)^2\sigma^2\}]$
$= exp[\{\theta_1 + \theta_3(\beta_0 + \beta_1 a^* + \beta_2^{'} c + \theta_2\sigma^2)\}(a - a^*) + 0.5\theta_3^2\sigma^2(a^2 - a^{*2})].$

If the regression models (3) and (4) are correctly specified and assumptions (i)-(iv) hold, the outcome Y is rare, and the error term for linear regression model (3) is normally distributed and has constant variance σ^2, then we could compute the natural indirect effects by:

$OR^{NIE} = exp[log\{\frac{P(Y_{aM_a}=1|c)/(1-P(Y_{aM_a}=1|c))}{P(Y_{aM_{a^*}}=1|c)/(1-P(Y_{aM_{a^*}}=1|c))}\}]$
$= exp[logit\{P(Y_{aM_a} = 1|c)\} - logit\{P(Y_{aM_{a^*}} = 1|c)\}]$
$\sim exp[\theta_0 + \theta_1 a + \theta_4^{'} c + (\theta_2 + \theta_3 a)(\beta_0 + \beta_1 a + \beta_2^{'} c) + \frac{1}{2}(\theta_2 + \theta_3 a)^2\sigma^2 - \{\theta_0 + \theta_1 a + \theta_4^{'} c + (\theta_2 + \theta_3 a)(\beta_0 + \beta_1 a^* + \beta_2^{'} c) + \frac{1}{2}(\theta_2 + \theta_3 a)^2\sigma^2\}]$
$= exp[(\theta_2\beta_1 + \theta_3\beta_1 a)(a - a^*)].$

If the regression models (3) and (4) are correctly specified and assumptions (i)-(iv) hold, the outcome Y is rare, and the error term for linear regression model (3) is normally distributed and has constant variance σ^2, then we could compute the total effects by:

$$OR^{TE} = exp[log\{\frac{P(Y_{aM_{a^*}}=1|c)/(1-P(Y_{aM_{a^*}}=1|c))}{P(Y_{a^*M_{a^*}}=1|c)/(1-P(Y_{a^*M_{a^*}}=1|c))}\}] \times exp[log\{\frac{P(Y_{aM_a}=1|c)/(1-P(Y_{aM_a}=1|c))}{P(Y_{aM_{a^*}}=1|c)/(1-P(Y_{aM_{a^*}}=1|c))}\}]$$

$$= E[Y_{a,M_{a^*}} - Y_{a^*,M_{a^*}}|C=c] \times E[Y_{a,M_a} - Y_{a^*,M_{a^*}}|C=c]$$
$$= exp[(\theta_1 + \theta_3\beta_0 + \theta_3\beta_1 a^* + \theta_3\beta_2^{'}c + \theta_2\beta_1 + \theta_3\beta_1 a + \theta_3\theta_2\sigma^2)(a-a^*) + 0.5\theta_3^2\sigma^2(a^2 - a^{*2})].$$

If the regression models (3) and (4) are correctly specified and assumptions (i)-(iv) hold then we can compute the proportion mediated by:

$$PM = \frac{log\{\frac{P(Y_{aM_a}=1|c)/(1-P(Y_{aM_a}=1|c))}{P(Y_{aM_{a^*}}=1|c)/(1-P(Y_{aM_{a^*}}=1|c))}\}}{log\{\frac{P(Y_{aM_{a^*}}=1|c)/(1-P(Y_{aM_{a^*}}=1|c))}{P(Y_{a^*M_{a^*}}=1|c)/(1-P(Y_{a^*M_{a^*}}=1|c))}\}+log\{\frac{P(Y_{aM_a}=1|c)/(1-P(Y_{aM_a}=1|c))}{P(Y_{aM_{a^*}}=1|c)/(1-P(Y_{aM_{a^*}}=1|c))}\}}$$
$$= \frac{(\theta_2\beta_1+\theta_3\beta_1 a)(a-a^*)}{(\theta_1+\theta_3\beta_0+\theta_3\beta_1 a^*+\theta_3\beta_2^{'}c+\theta_2\beta_1+\theta_3\beta_1 a+\theta_3\theta_2\sigma^2)(a-a^*)+0.5\theta_3^2\sigma^2(a^2-a^{*2})}.$$

These expressions apply also if the outcome is not rare and log-linear rather than logistic models are fit to the outcome model; the direct and indirect effect will have now an interpretation on the risk ratio scale rather than on the odds ratio scale.
These expressions apply also if the outcome is a count variable. In particular if $Y \sim Poi(\lambda)$ for $\lambda = \exp\{\theta_0 + \theta_1 a + \theta_2 m + \theta_3 a * m + \theta_4^{'}c\}$) the outcome regression can be defined as:

$$log\{E(Y|A=a, M=m, C=c)\} = \theta_0 + \theta_1 a + \theta_2 m + \theta_3 a * m + \theta_4^{'}c$$

The natural direct effect for binary outcome on the risk ratio scale coincides with the natural direct effect for poisson count outcome since:

$$RR^{NDE} = exp[log\{\frac{E(Y_{aM_{a^*}}|c)}{E(Y_{a^*M_{a^*}}|c)}\}]$$

The same argument holds for the natural indirect effect. Finally, the argument can be extended to the case in which the count outcome is modeled with a negative binomial distribution. This is the case since the negative binomial distribution can be represented as an over-dispersed poisson and the mean of the two models coincide.

Standard errors

We now consider standard errors for the controlled direct effect and natural direct and indirect effect odds ratios. Suppose that model (4) has been fit using standard logistic regression software and that model (3) has been fit using standard linear regression software. Suppose furthermore that the resulting estimates $\hat{\beta}$ of $\beta = (\beta_0, \beta_1, \beta_2^{'})'$, $\hat{\theta}$ of $\theta = (\theta_0, \theta_1, \theta_2, \theta_3, \theta_4^{'})'$ and $\hat{\sigma}^2$ of σ have covariance matrices Σ_β and Σ_θ. Then the covariance matrix of $(\hat{\beta}', \hat{\theta}', \hat{\sigma}^2)$ is

$$\Sigma = \begin{bmatrix} \Sigma_\beta & 0 & 0 \\ 0 & \Sigma_\theta & 0 \\ 0 & 0 & \Sigma_{\sigma^2} \end{bmatrix}$$

Standard errors of the controlled and natural direct and indirect effects can be obtained (using the delta method) as

$$\sqrt{\Gamma\Sigma\Gamma'}|a - a^*|$$

with $\Gamma = (0, 0, 0^{'}, 0, 1, 0, m, 0^{'}, 0)$ for the log of controlled direct effect odds ratio, $\Gamma = (\theta_3, \theta_3 a^*, \theta_3 c^{'}, 0, 1, \theta_3\sigma^2, \beta_0 + \beta_1 a^* + \beta_2^{'} c + \theta_2\sigma^2 + \theta_3\sigma^2(a + a^*), 0^{'}, \theta_2\theta_3 + 0.5\theta_3^2(a + a^*))$ for the log pure natural direct effect odds ratio (same expression holds for the total natural direct effect upon substituting a and a^*), $\Gamma = (0, \theta_2 + \theta_3 a, 0^{'}, 0, 0, \beta_1, \beta_1 a, 0^{'}, 0)$ for the log of total natural indirect effect (the same expression holds for the pure natural indirect effect upon substituting a and a^*), $\Gamma = (\theta_3, \theta_3(a + a^*) + \theta_2, \theta_3 c^{'}, 0, 1, \theta_3\sigma^2 + \beta_1, \beta_0 + \beta_1(a + a^*) + \beta_2^{'} c + \theta_2\sigma^2 + \theta_3\sigma^2(a^2 - a^{*2}), 0^{'}, 0.5\theta_3^2(a^2 - a^{*2}))$ for the logarithm of the total effect. Standard errors for the proportion mediated can be obtained (using the delta method) as

$$\sqrt{\Gamma\Sigma\Gamma'}$$

where $\Gamma = (d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8, d_9)$.

Let

$$A = (\theta_2\beta_1 + \theta_3\beta_1 a)(a - a^*)$$

$$B = [\{\theta_1 + \theta_3(\beta_0 + \beta_1(a + a^*) + \beta_2' c + \theta_2\sigma^2) + \beta_1\theta_2\}(a - a^*) + 0.5\theta_3^2\sigma^2(a^2 - a^{*2})]$$

$$d_1 = -\frac{\theta_3(a - a^*)A}{B^2}$$

$$d_2 = \frac{(\theta_2 + \theta_3 a)(a - a^*)B - (\theta_3(a + a^*) + \theta_2)(a - a^*)A}{B^2}$$

$$d_3 = -\frac{\theta_3' c(a - a^*)A}{B^2}$$

$$d_4 = 0$$

$$d_5 = -\frac{A(a - a^*)}{B^2}$$

$$d_6 = \frac{\beta_1(a - a^*)B - (\theta_2\sigma^2 + \beta_1)(a - a^*)A}{B^2}$$

$$d_7 = \frac{\beta_1 a(a-a^*)B + (\beta_0 + \beta_1(a+a^*) + \beta_2' c + \theta_2\sigma^2)(a-a^*) - (\theta_2\sigma^2)(a-a^*)A}{B^2}$$

$$d_8 = 0'$$

$$d_9 = -\frac{[\theta_3\theta_2(a-a^*) + 0.5\theta_3^2(a^2 - a^{*2})]A}{B^2}$$

3 Binary Mediator and Continuous Outcome

Effects using regression

Suppose that the outcome is continuous, the mediator is binary and that the following models fit the observed data:

$$logit\{P(M=1|A=a, C=c)\} = \beta_0 + \beta_1 a + \beta_2' c \tag{5}$$

$$E(Y|A=a, M=m, C=c) = \theta_0 + \theta_1 a + \theta_2 m + \theta_3 a * m + \theta_4' c \tag{6}$$

In particular, if the regression models (5) and (6) are correctly specified and assumptions (i) and (ii) hold then we could compute the average controlled direct effect as in section 1

If the regression models (5) and (6) are correctly specified and assumptions (i)-(iv) hold then we could compute the average natural direct effects by:

$NDE = E[Y_{aM_{a^*}} - Y_{a^*M_{a^*}}|C=c]$
$= \sum_m \{E[Y|C=c, A=a, M=m] - E[Y|C=c, A=a^*, M=m]\} \times P(M=m|C=c, A=a^*)$

$= \sum_m \{(\theta_0 + \theta_1 a + \theta_2 m + \theta_3 am + \theta_4' c) - (\theta_0 + \theta_1 a^* + \theta_2 m + \theta_3 a^* m + \theta_4' c)\} \times P(M = m|C = c, A = a^*)$

$= \sum_m \{(\theta_1 a + \theta_2 m + \theta_3 am) - (\theta_1 a^* + \theta_2 m + \theta_3 a^* m)\} \times P(M = m|C = c, A = a^*)$

$= \{\theta_1(a - a^*)\} + \{\theta_3(a - a^*)\} \frac{exp[\beta_0 + \beta_1 a^* + \beta_2' c]}{1 + exp[\beta_0 + \beta_1 a^* + \beta_2' c]}$.

If the regression models (5) and (6) are correctly specified and assumptions (i)-(iv) hold then we could compute the average natural indirect effects by:

$NIE = E[Y_{aM_a} - Y_{aM_{a^*}}|C = c]$

$= \sum_m E[Y|C = c, A = a, M = m] \times P(M = m|C = c, A = a) - \sum_m E[Y|C = c, A = a, M = m] \times P(M = m|C = c, A = a^*)$

$= \sum_m (\theta_0 + \theta_1 a + \theta_2 m + \theta_3 am + \theta_4' c) \times P(M = m|C = c, A = a) - \sum_m (\theta_0 + \theta_1 a + \theta_2 m + \theta_3 am + \theta_4' c) \times P(M = m|C = c, A = a^*)$

$= (\theta_2 + \theta_3 a)\{E[M|A = a, C = c] - E[M|A = a^*, C = c]\}$

$=(\theta_2 + \theta_3 a)\{\frac{exp[\beta_0 + \beta_1 a + \beta_2' c]}{1 + exp[\beta_0 + \beta_1 a + \beta_2' c]} - \frac{exp[\beta_0 + \beta_1 a^* + \beta_2' c]}{1 + exp[\beta_0 + \beta_1 a^* + \beta_2' c]}\}$.

If the regression models (5) and (6) are correctly specified and assumptions (i)-(iv) hold then we could compute the total effect by:

$TE = E[Y_a - Y_{a^*}|C = c]$

$= E[Y_{aM_{a^*}} - Y_{a^*M_{a^*}}|C = c] + E[Y_{aM_a} - Y_{a^*M_{a^*}}|C = c]$

$= \{\theta_1(a - a^*)\} + \{\theta_3(a - a^*)\} \frac{exp[\beta_0 + \beta_1 a^* + \beta_2' c]}{1 + exp[\beta_0 + \beta_1 a^* + \beta_2' c]} + (\theta_2 + \theta_3 a)\{\frac{exp[\beta_0 + \beta_1 a + \beta_2' c]}{1 + exp[\beta_0 + \beta_1 a + \beta_2' c]} - \frac{exp[\beta_0 + \beta_1 a^* + \beta_2' c]}{1 + exp[\beta_0 + \beta_1 a^* + \beta_2' c]}\}$.

If the regression models (5) and (6) are correctly specified and assumptions (i)-(iv) hold then we could compute the proportion mediated by:

$PM = \frac{E[Y_{aM_a} - Y_{a^*M_{a^*}}|C=c]}{E[Y_a - Y_{a^*}|C=c]}$

$$= \frac{(\theta_2+\theta_3 a)\{\frac{exp[\beta_0+\beta_1 a+\beta_2^{'}c]}{1+exp[\beta_0+\beta_1 a+\beta_2^{'}c]}-\frac{exp[\beta_0+\beta_1 a^*+\beta_2^{'}c]}{1+exp[\beta_0+\beta_1 a^*+\beta_2^{'}c]}\}}{(\theta_2+\theta_3 a)\{\frac{exp[\beta_0+\beta_1 a+\beta_2^{'}c]}{1+exp[\beta_0+\beta_1 a+\beta_2^{'}c]}-\frac{exp[\beta_0+\beta_1 a^*+\beta_2^{'}c]}{1+exp[\beta_0+\beta_1 a^*+\beta_2^{'}c]}\}+\{\theta_1(a-a^*)\}+\{\theta_3(a-a^*)\}\frac{exp[\beta_0+\beta_1 a^*+\beta_2^{'}c]}{1+exp[\beta_0+\beta_1 a^*+\beta_2^{'}c]}}.$$

Standard errors

Suppose that model (6) have been fit using standard linear regression software and that model (5) have been fit using standard logistic regression. The resulting estimates are $\hat{\beta}$ of $\beta = (\beta_0, \beta_1, \beta_2^{'})^{'}$and $\hat{\theta}$ of $\theta = (\theta_0, \theta_1, \theta_2, \theta_3, \theta_4^{'})^{'}$ have covariance matrices Σ_β and Σ_θ. Then the covariance matrix of $(\hat{\beta}^{'}, \hat{\theta}^{'})$ is

$$\Sigma = \begin{bmatrix} \Sigma_\beta & 0 \\ 0 & \Sigma_\theta \end{bmatrix}$$

Standard errors of the controlled and natural direct can be obtained (using the delta method) as

$$\sqrt{\Gamma\Sigma\Gamma'}|a - a^*|$$

with $\Gamma = (0, 0, 0^{'}, 0, 1, 0, m, 0^{'})$ for the controlled direct effect, $\Gamma = (d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8)$ for the pure natural direct effect (same expression holds for the total natural direct effect upon substituting a and a^*), where

$$d_1 = \frac{\theta_3 exp[[\beta_0 + \beta_1 a^{*'}_{\beta 2} c](1 + exp[\beta_0 + \beta_1 a^* \beta_2' c]) - \theta_3\{exp[\beta_0 + \beta_1 a^* \beta_2' c]\}^2}{(1 + exp[\beta_0 + \beta_1 a^* \beta_2' c])^2}$$

$$d_2 = \frac{\theta_3 a^* exp[[\beta_0 + \beta_1 a^{*'}_{\beta 2} c](1 + exp[\beta_0 + \beta_1 a^* \beta_2' c]) - \{exp[\beta_0 + \beta_1 a^* \beta_2' c]\}^2}{(1 + exp[\beta_0 + \beta_1 a^* \beta_2' c])^2}$$

$$d_3 = \frac{\theta_3 c^{'} exp[[\beta_0 + \beta_1 a^{*'}_{\beta 2} c](1 + exp[\beta_0 + \beta_1 a^* \beta_2' c]) - \{exp[\beta_0 + \beta_1 a^* \beta_2' c]\}^2}{(1 + exp[\beta_0 + \beta_1 a^* \beta_2' c])^2}$$

$$d_4 = 0$$

$$d_5 = 1$$

$$d_6 = 0$$

$$d_7 = \frac{exp[\beta_0 + \beta_1 a^* \beta_2^{'} c]}{1 + exp[\beta_0 + \beta_1 a^* \beta_2^{'} c]}$$

$$d_8 = 0^{'}$$

Standard errors of the natural indirect can be obtained (using the delta method) as

$$\sqrt{\Gamma \Sigma \Gamma'}$$

For the natural indirect effect (the same expression holds for the pure natural indirect effect upon substituting a and a^*) let

$$A = \frac{exp[\beta_0 + \beta_1 a + \beta_2^{'} c]\{1 + exp[\beta_0 + \beta_1 a + \beta_2^{'} c]\} - \{exp[\beta_0 + \beta_1 a + \beta_2^{'} c]\}^2}{\{1 + exp[\beta_0 + \beta_1 a + \beta_2^{'} c]\}^2}$$

$$B = \frac{exp[\beta_0 + \beta_1 a^* + \beta_2^{'} c]\{1 + exp[\beta_0 + \beta_1 a^* + \beta_2^{'} c]\} - \{exp[\beta_0 + \beta_1 a^* + \beta_2^{'} c]\}^2}{\{1 + exp[\beta_0 + \beta_1 a^* + \beta_2^{'} c]\}^2}$$

$$K = \frac{exp[\beta_0 + \beta_1 a + \beta_2^{'} c]}{\{1 + exp[\beta_0 + \beta_1 a + \beta_2^{'} c]\}}$$

$$D = \frac{exp[\beta_0 + \beta_1 a^* + \beta_2^{'} c]}{\{1 + exp[\beta_0 + \beta_1 a^* + \beta_2^{'} c]\}}$$

and

$\Gamma = (d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8)$, where

$$d_1 = \{\theta_2 + \theta_3 a\}[A - B]$$

$$d_2 = \{\theta_2 + \theta_3 a\}[aA - a^* B]$$

$$d_3 = \{\theta_2 + \theta_3 a\}c^{'}[A - B]$$

$$d_4 = 0$$

$$d_5 = 0$$

$$d_6 = K - D$$

$$d_7 = a[K - D]$$

$$d_8 = 0^{'}$$

Standard errors of the controlled and total effect and percentage mediated can be obtained (using the delta method) as

$$\sqrt{\Gamma\Sigma\Gamma'}$$

let

$$A = \frac{exp[\beta_0 + \beta_1 a + \beta_2^{'} c]\{1 + exp[\beta_0 + \beta_1 a + \beta_2^{'} c]\} - \{exp[\beta_0 + \beta_1 a + \beta_2^{'} c]\}^2}{\{1 + exp[\beta_0 + \beta_1 a + \beta_2^{'} c]\}^2}$$

$$B = \frac{exp[\beta_0 + \beta_1 a^* + \beta_2^{'} c]\{1 + exp[\beta_0 + \beta_1 a^* + \beta_2^{'} c]\} - \{exp[\beta_0 + \beta_1 a^* + \beta_2^{'} c]\}^2}{\{1 + exp[\beta_0 + \beta_1 a^* + \beta_2^{'} c]\}^2}$$

$$K = \frac{exp[\beta_0 + \beta_1 a + \beta_2^{'} c]}{\{1 + exp[\beta_0 + \beta_1 a + \beta_2^{'} c]\}}$$

$$D = \frac{exp[\beta_0 + \beta_1 a^* + \beta_2^{'} c]}{\{1 + exp[\beta_0 + \beta_1 a^* + \beta_2^{'} c]\}}$$

for the total effect $\Gamma = (d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8)$, where

$$d_1 = \theta_3(a - a^*)B + (\theta_2 + \theta_3 a)(A - B)$$

$$d_2 = a^*\theta_3(a - a^*)B + (\theta_2 + \theta_3 a)(aA - a^*B)$$

$$d_3 = c^{'}\theta_3(a - a^*)B + (\theta_2 + \theta_3 a)(A - B)$$

$$d_4 = 0$$

$$d_5 = a - a^*$$

$$d_6 = K - D$$

$$d_7 = (a - a^*)D + a[K - D]$$

$$d_8 = 0^{'}$$

and for the proportion mediated $\Gamma = (d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8)$ where

$$d_1 = \frac{[(\theta_2+\theta_3 a)(A-B)]\{(\theta_2+\theta_3 a)(K-D)+(a-a^*)[\theta_1+\theta_3 D]\}-\{[(\theta_2+\theta_3 a)(A-B)]+(a-a^*)\theta_3 B\}(\theta_2+\theta_3 a)[K-D]}{\{(\theta_2+\theta_3 a)(K-D)+(a-a^*)[\theta_1+\theta_3 D]\}^2}$$
$$d_2 = \frac{[(\theta_2+\theta_3 a)(aA-a^* B)]\{(\theta_2+\theta_3 a)(K-D)+(a-a^*)[\theta_1+\theta_3 D]\}-\{[(\theta_2+\theta_3 a)(aA-a^* B)]+a^*(a-a^*)\theta_3 B\}(\theta_2+\theta_3 a)[K-D]}{\{(\theta_2+\theta_3 a)(K-D)+(a-a^*)[\theta_1+\theta_3 D]\}^2}$$
$$d_3 = \frac{[(\theta_2+\theta_3 a)c^{'}(A-B)]\{(\theta_2+\theta_3 a)(K-D)+(a-a^*)[\theta_1+\theta_3 D]\}-c^{'}\{[(\theta_2+\theta_3 a)(A-B)]+(a-a^*)\theta_3 B\}(\theta_2+\theta_3 a)[K-D]}{\{(\theta_2+\theta_3 a)(K-D)+(a-a^*)[\theta_1+\theta_3 D]\}^2}$$

$$d_4 = 0$$

$$d_5 = \frac{(a - a^*)(\theta_2 + \theta_3 a)[K - D]}{\{(\theta_2 + \theta_3 a)(K - D) + (a - a^*)[\theta_1 + \theta_3 D]\}^2}$$

$$d_6 = \frac{a[K-D]\{(\theta_2+\theta_3 a)(K-D)+(a-a^*)[\theta_1+\theta_3 D]\}-[K-D]\{(\theta_2+\theta_3 a)(K-D)+(a-a^*)[\theta_1+\theta_3 D]\}}{\{(\theta_2+\theta_3 a)(K-D)+(a-a^*)[\theta_1+\theta_3 D]\}^2}$$
$$d_7 = \frac{[K-D]\{(\theta_2+\theta_3 a)(K-D)+(a-a^*)[\theta_1+\theta_3 D]\}-\{a[K-D]+(a-a^* D)\}\{(\theta_2+\theta_3 a)(K-D)+(a-a^*)[\theta_1+\theta_3 D]\}}{\{(\theta_2+\theta_3 a)(K-D)+(a-a^*)[\theta_1+\theta_3 D]\}^2}$$

$$d_8 = 0^{'}.$$

4 Binary Mediator and Binary Outcome

Effects using regression

Suppose that both the outcome and the mediator are binary and that the following models fit the observed data:

$$logit\{P(M=1|A=a,C=c)\}=\beta_0+\beta_1 a+\beta_2^{'}c \tag{7}$$

$$logit\{P(Y=1|A=a,M=m,C=c)\}=\theta_0+\theta_1 a+\theta_2 m+\theta_3 a*m+\theta_4^{'}c \tag{8}$$

If the regression models (7) and (8) are correctly specified and assumptions (i) and (ii) hold then we can compute the controlled direct effect odds ratio as the case in which the mediator is continuous and the outcome is binary.

If the regression models (7) and (8) are correctly specified and assumptions (i)-(iv) hold and the outcome Y is rare, then we could compute the average natural direct effects by:

$$OR^{NDE}=exp[log\{\frac{P(Y_{aM_{a^*}}=1|c)/(1-P(Y_{aM_{a^*}}=1|c))}{P(Y_{a^*M_{a^*}}=1|c)/(1-P(Y_{a^*M_{a^*}}=1|c))}\}]$$
$$=exp[logit\{P(Y_{aM_{a^*}}=1|c)\}-logit\{P(Y_{a^*M_{a^*}}=1|c)\}]$$
$$\sim exp[log\{\frac{exp(\theta_0+\theta_1 a+\theta_4^{'}c)+exp(\theta_0+\theta_1 a+\theta_4^{'}c+\theta_2+\theta_3 a+\beta_0+\beta_1 a^*+\beta_2^{'}c)}{1+exp[\beta_0+\beta_1 a^*+\beta_2^{'}C]}\}-log\{\frac{exp(\theta_0+\theta_1 a^*+\theta_4^{'}c)+exp(\theta_0+\theta_1 a^*+\theta_4^{'}c+\theta_2+\theta_3 a^*+\beta_0+\beta_1 a^*+\beta_2^{'}c)}{1+exp[\beta_0+\beta_1 a^*+\beta_2^{'}c]}\}]$$
$$=\{\frac{exp[\theta_0+\theta_1 a+\theta_4^{'}c]+exp[\theta_0+\theta_1 a+\theta_4^{'}c+\theta_2+\theta_3 a+\beta_0+\beta_1 a^*+\beta_2^{'}c]}{exp[\theta_0+\theta_1 a^*+\theta_4^{'}c]+exp(\theta_0+\theta_1 a^*+\theta_4^{'}c+\theta_2+\theta_3 a^*+\beta_0+\beta_1 a^*+\beta_2^{'}c)}\}$$

$$=\{\frac{exp[\theta_1 a](1+exp[\theta_2+\theta_3 a+\beta_0+\beta_1 a^*+\beta_2^{'}c])}{exp[\theta_1 a^*](1+exp[\theta_2+\theta_3 a^*+\beta_0+\beta_1 a^*+\beta_2^{'}c])}\}.$$

If the regression models (7) and (8) are correctly specified and assumptions (i)-(iv) hold and the outcome Y is rare, then we could compute the average natural indirect effects by:

$$OR^{NIE}=exp[log\{\frac{P(Y_{aM_a}=1|c)/(1-P(Y_{aM_a}=1|c))}{P(Y_{aM_{a^*}}=1|c)/(1-P(Y_{aM_{a^*}}=1|c))}\}]$$
$$=exp[logit\{P(Y_{aM_a}=1|c)\}-logit\{P(Y_{aM_{a^*}}=1|c)\}]$$

$$\sim exp[log\{\frac{exp(\theta_0+\theta_1a+\theta_4^{'}c)+exp(\theta_0+\theta_1a+\theta_4^{'}c+\theta_2+\theta_3a+\beta_0+\beta_1a+\beta_2^{'}c)}{1+exp[\beta_0+\beta_1a+\beta_2^{'}c]}\}-log\{\frac{exp(\theta_0+\theta_1a+\theta_4^{'}c)+exp(\theta_0+\theta_1a+\theta_4^{'}c+\theta_2+\theta_3a+\beta_0+\beta_1a^*+\beta_2^{'}c)}{1+exp[\beta_0+\beta_1a^*+\beta_2^{'}c]}\}]$$
$$=\frac{[1+exp(\beta_0+\beta_1a^*+\beta_2^{'}c)][1+exp(\theta_2+\theta_3a+\beta_0+\beta_1a+\beta_2^{'}c)]}{[1+exp(\beta_0+\beta_1a+\beta_2^{'}c)][1+exp(\theta_2+\theta_3a+\beta_0+\beta_1a^*+\beta_2^{'}c)]}.$$

If the regression models (7) and (8) are correctly specified and assumptions (i)-(iv) hold, the outcome Y is rare, then we could compute the total effects by:

$$OR^{TE}=exp[log\{\frac{P(Y_{aM_{a^*}}=1|c)/(1-P(Y_{aM_{a^*}}=1|c))}{P(Y_{a^*M_{a^*}}=1|c)/(1-P(Y_{a^*M_{a^*}}=1|c))}\}]\times exp[log\{\frac{P(Y_{aM_a}=1|c)/(1-P(Y_{aM_a}=1|c))}{P(Y_{aM_{a^*}}=1|c)/(1-P(Y_{aM_{a^*}}=1|c))}\}]$$

$$=\{\frac{exp[\theta_1a](1+exp[\theta_2+\theta_3a+\beta_0+\beta_1a^*+\beta_2^{'}c])}{exp[\theta_1a^*](1+exp[\theta_2+\theta_3a^*+\beta_0+\beta_1a^*+\beta_2^{'}c])}\}\times\{\frac{[1+exp(\beta_0+\beta_1a^*+\beta_2^{'}c)][1+exp(\theta_2+\theta_3a+\beta_0+\beta_1a+\beta_2^{'}c)]}{[1+exp(\beta_0+\beta_1a+\beta_2^{'}c)][1+exp(\theta_2+\theta_3a+\beta_0+\beta_1a^*+\beta_2^{'}c)]}\}.$$

If the regression models (7) and (8) are correctly specified and assumptions (i)-(iv) hold then we can compute the proportion mediated by:

$$PM=\frac{log\{\frac{P(Y_{aM_a}=1|c)/(1-P(Y_{aM_a}=1|c))}{P(Y_{aM_{a^*}}=1|c)/(1-P(Y_{aM_{a^*}}=1|c))}\}}{log\{\frac{P(Y_{aM_{a^*}}=1|c)/(1-P(Y_{aM_{a^*}}=1|c))}{P(Y_{a^*M_{a^*}}=1|c)/(1-P(Y_{a^*M_{a^*}}=1|c))}\}+log\{\frac{P(Y_{aM_a}=1|c)/(1-P(Y_{aM_a}=1|c))}{P(Y_{aM_{a^*}}=1|c)/(1-P(Y_{aM_{a^*}}=1|c))}\}}$$
$$=\frac{log\frac{[1+exp(\beta_0+\beta_1a^*+\beta_2^{'}c)][1+exp(\theta_2+\theta_3a+\beta_0+\beta_1a+\beta_2^{'}c)]}{[1+exp(\beta_0+\beta_1a+\beta_2^{'}c)][1+exp(\theta_2+\theta_3a+\beta_0+\beta_1a^*+\beta_2^{'}c)]}}{log\{\frac{exp[\theta_1a](1+exp[\theta_2+\theta_3a+\beta_0+\beta_1a^*+\beta_2^{'}c])}{exp[\theta_1a^*](1+exp[\theta_2+\theta_3a^*+\beta_0+\beta_1a^*+\beta_2^{'}c])}\}\times\{\frac{[1+exp(\beta_0+\beta_1a^*+\beta_2^{'}c)][1+exp(\theta_2+\theta_3a+\beta_0+\beta_1a+\beta_2^{'}c)]}{[1+exp(\beta_0+\beta_1a+\beta_2^{'}c)][1+exp(\theta_2+\theta_3a+\beta_0+\beta_1a^*+\beta_2^{'}c)]}\}}.$$

These expressions apply also if the outcome is not rare and log-linear rather than logistic models are fit to the outcome model; the direct and indirect effect will have now an interpretation on the risk ratio scale rather than on the odds ratio scale.
These expressions apply also if the outcome is a count variable. In particular if $Y\sim Poi(\lambda)$ for $\lambda=\exp\{\theta_0+\theta_1a+\theta_2m+\theta_3a*m+\theta_4^{'}c\})$ the outcome regression can be defined as:

$$log\{E(Y|A=a,M=m,C=c)\}=\theta_0+\theta_1a+\theta_2m+\theta_3a*m+\theta_4^{'}c$$

The natural direct effect for binary outcome on the risk ratio scale coincides with the natural direct effect for poisson count outcome since:

$$RR^{NDE} = exp[log\{\frac{E(Y_{aM_{a^*}}|c)}{E(Y_{a^*M_{a^*}}|c)}\}]$$

The same argument holds for the natural indirect effect. Finally, the argument can be extended to the case in which the count outcome is modeled with a negative binomial distribution. This is the case since the negative binomial distribution can be represented as an over-dispersed poisson and the mean of the two models coincide.

Standard Errors:

Suppose that model (7) and (8) have been fit using standard logistic regression software and that the resulting estimates $\hat{\beta}$ of $\beta = (\beta_0, \beta_1, \beta_2^{'})^{'}$ and $\hat{\theta}$ of $\theta = (\theta_0, \theta_1, \theta_2, \theta_3, \theta_4^{'})^{'}$ have covariance matrices Σ_β and Σ_θ. Then the covariance matrix of $(\hat{\beta}^{'}, \hat{\theta}^{'})$ is

$$\Sigma = \begin{bmatrix} \Sigma_\beta & 0 \\ 0 & \Sigma_\theta \end{bmatrix}$$

Standard errors of the controlled and natural direct and indirect effects can be obtained (using the delta method) as

$$\sqrt{\Gamma\Sigma\Gamma^{'}}$$

with $\Gamma = (0, 0, 0^{'}, 0, (a - a^*), 0, m(a - a^*), 0^{'})$ for the controlled direct effect, $\Gamma = (d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8)$ for the logarithm of the pure natural direct effect (same expression holds for the logarithm of the total natural direct effect upon substituting a and a^*), where let

$$A = \frac{exp[\theta_2 + \theta_3 a + \beta_0 + \beta_1 a^* + \beta_2^{'} c]}{\{1 + exp[\theta_2 + \theta_3 a + \beta_0 + \beta_1 a^* + \beta_2^{'} c]\}}$$

$$B = \frac{exp[\theta_2 + \theta_3 a^* + \beta_0 + \beta_1 a^* + \beta_2' c]}{\{1 + exp[\theta_2 + \theta_3 a^* + \beta_0 + \beta_1 a^* + \beta_2' c]\}}$$

and

$$d_1 = A - B$$

$$d_2 = a^*(A - B)$$

$$d_3 = c^{'}(A - B)$$

$$d_4 = 0$$

$$d_5 = (a - a^*)$$

$$d_6 = A - B$$

$$d_7 = aA - a^* B$$

$$d_8 = 0^{'}$$

for the logarithm of the natural indirect effect (the same expression holds for the pure natural indirect effect upon substituting a and a^*) let

$$A = \frac{exp[\theta_2 + \theta_3 a + \beta_0 + \beta_1 a + \beta_2' c]}{\{1 + exp[\theta_2 + \theta_3 a + \beta_0 + \beta_1 a + \beta_2' c]\}}$$

$$B = \frac{exp[\theta_2 + \theta_3 a + \beta_0 + \beta_1 a^* + \beta_2' c]}{\{1 + exp[\theta_2 + \theta_3 a + \beta_0 + \beta_1 a^* + \beta_2' c]\}}$$

$$K = \frac{exp[\beta_0 + \beta_1 a + \beta_2' c]}{\{1 + exp[\beta_0 + \beta_1 a + \beta_2' c]\}}$$

$$D = \frac{exp[\beta_0 + \beta_1 a^* + \beta_2' c]}{\{1 + exp[\beta_0 + \beta_1 a^* + \beta_2' c]\}}$$

and

$\Gamma = (d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8)$ where

$$d_1 = (D + A) - (K + B)$$

$$d_2 = a^*[D - B] + a[A - K]$$

$$d_3 = c^{'}[(D + A) - (K + B)]$$

$$d_4 = 0$$

$$d_5 = 0$$

$$d_6 = A - B$$

$$d_7 = a[A - B]$$

$$d_8 = 0^{'}$$

Standard errors of the logarithm of the total effect and percentage mediated can be obtained (using the delta method) as

$$\sqrt{\Gamma\Sigma\Gamma'}$$

Let $d_i(log(pnde))$ and $d_i(log(tnie))$ for $i = 1, ..., 8$, the gamma elements derived for the logarithm of the pure natural direct effect and the total natural indirect effect respectively.

For the total effect $\Gamma = (d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8)$,where

$$d_1 = d_1(log(pnde)) + d_1(log(tnie))$$

$$d_2 = d_2(log(pnde)) + d_2(log(tnie))$$

$$d_3 = d_3(log(pnde)) + d_3(log(tnie))$$

$$d_4 = d_4(log(pnde)) + d_4(log(tnie))$$

$$d_5 = d_5(log(pnde)) + d_5(log(tnie))$$

$$d_6 = d_6(log(pnde)) + d_6(log(tnie))$$

$$d_7 = d_7(log(pnde)) + d_7(log(tnie))$$

$$d_8 = d_8(log(pnde)) + d_8(log(tnie))$$

and for the proportion mediated $\Gamma = (d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8)$. Let

$$d_1 = \frac{d_1(log(tnie)) * log(te) - d_1(log(te)) * log(tnie)}{log(te)^2}$$

$$d_2 = \frac{d_2(log(tnie)) * log(te) - d_2(log(te)) * log(tnie)}{log(te)^2}$$

$$d_3 = \frac{d_3(log(tnie)) * log(te) - d_3(log(te)) * log(tnie)}{log(te)^2}$$

$$d_4 = \frac{d_4(log(tnie)) * log(te) - d_4(log(te)) * log(tnie)}{log(te)^2}$$

$$d_5 = \frac{d_5(log(tnie)) * log(te) - d_5(log(te)) * log(tnie)}{log(te)^2}$$

$$d_6 = \frac{d_6(log(tnie)) * log(te) - d_6(log(te)) * log(tnie)}{log(te)^2}$$

$$d_7 = \frac{d_7(log(tnie)) * log(te) - d_7(log(te)) * log(tnie)}{log(te)^2}$$

$$d_8 = \frac{d_8(log(tnie)) * log(te) - d_8(log(te)) * log(tnie)}{log(te)^2}.$$